# Measuring Collaborative Cognition


Nancy J. Cooke

Arizona State University


CKM Workshop

January 14-16, 2003

# Report Documentation Page

*Form Approved*
*OMB No. 0704-0188*



| | | |
|---|---|---|
| 1. REPORT DATE<br>**JAN 2003** | 2. REPORT TYPE | 3. DATES COVERED<br>**00-00-2003 to 00-00-2003** |

| | |
|---|---|
| 4. TITLE AND SUBTITLE<br>**Measuring Collaborative Cognition** | 5a. CONTRACT NUMBER<br>5b. GRANT NUMBER<br>5c. PROGRAM ELEMENT NUMBER |
| 6. AUTHOR(S) | 5d. PROJECT NUMBER<br>5e. TASK NUMBER<br>5f. WORK UNIT NUMBER |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)<br>**Arizona State University,Tempe,AZ,85287** | 8. PERFORMING ORGANIZATION REPORT NUMBER |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) | 10. SPONSOR/MONITOR’S ACRONYM(S)<br>11. SPONSOR/MONITOR’S REPORT NUMBER(S) |

12. DISTRIBUTION/AVAILABILITY STATEMENT
**Approved for public release; distribution unlimited**

13. SUPPLEMENTARY NOTES
**Collaboration and Knowledge Management (CKM) Workshop, 14-16 Jan 2003, College Park, MD. U.S. Government or Federal Rights License**

14. ABSTRACT

15. SUBJECT TERMS

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT<br>**unclassified** | b. ABSTRACT<br>**unclassified** | c. THIS PAGE<br>**unclassified** | **Same as Report (SAR)** | **50** | |

**Standard Form 298 (Rev. 8-98)**
Prescribed by ANSI Std Z39-18

# Acknowledgements

- **NMSU Faculty**: Peter Foltz
- **NMSU Post Doc:** Brian Bell
- **NMSU Graduate Students:** Janie DeJoode, Jamie Gorman, Preston Kiekel, Rebecca Keith, Melanie Martin, Harry Pedersen
- **US Positioning, LLC:** Steven Shope
- **UCF:** Eduardo Salas, Clint Bowers
- **Sponsors**: Air Force Office of Scientific Research, Office of Naval Research, NASA Ames Research Center, Army Research Laboratory

# Overview

- What is Collaborative Cognition?
- A Focus on Measurement
- Assessing Collaborative Performance & Cognition
- Toward Diagnosis of Collaborative Performance
- Conclusions

# What is Collaborative Cognition?

# Collaborative Cognition in Practice

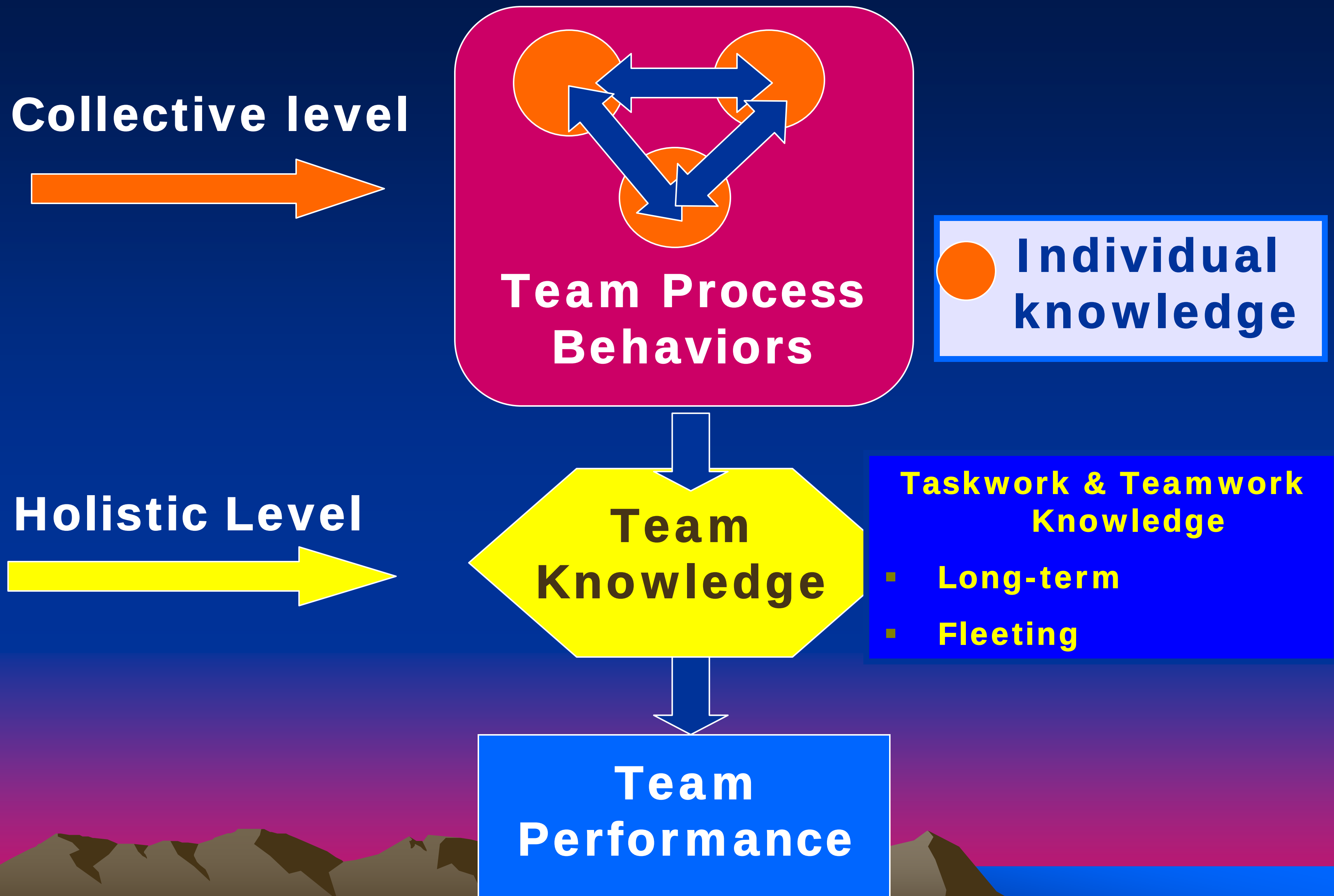
# Experimental Context

## CERTT (Cognitive Engineering Research on Team Tasks) Lab

**A Synthetic Task Environment for the Study of Collaborative Cognition**

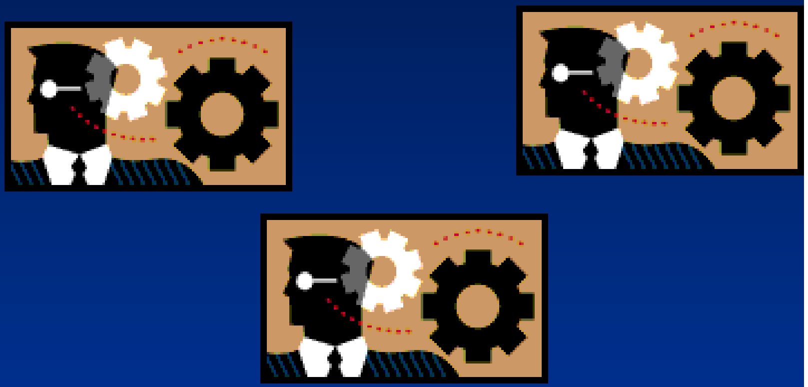

**Five Participant Consoles** **Experimenter Console**

# Collaborative Cognition Framework

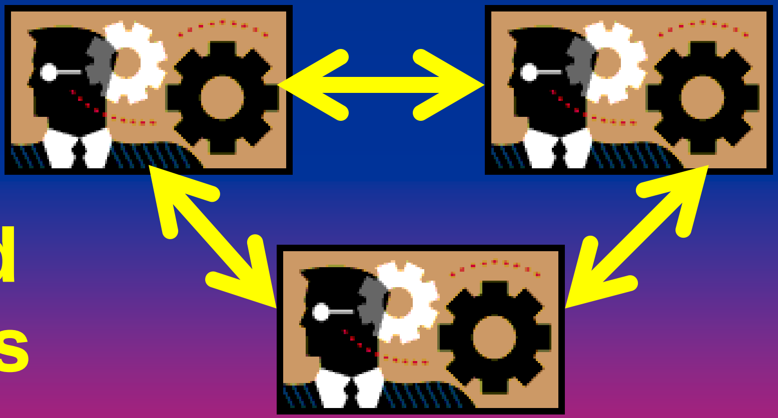


# Defining Collaborative Cognition

- It is more than the sum of the cognition of individual team members.



- It emerges from the interplay of the individual cognition of each team member and team process behaviors



# A Focus on Measurement

# Why Measurement?

- Assessment of collaborative performance or effectiveness (criterion or dependent measures) often taken for granted.
- Outcome measures of collaborative performance do not reveal why performance is effective or ineffective.
- Process measures of collaborative behavior are often subjective and lack reliability and validity.

# Why Measurement? (continued)

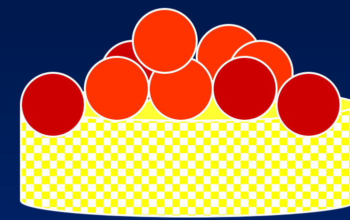
- Collaborative cognition is assumed to contribute to collaborative performance, and especially for growing number of cognitive tasks.
- Understanding the team cognition behind team performance should inform interventions (design, training, selection) to improve that performance.

# Measurement Limitations

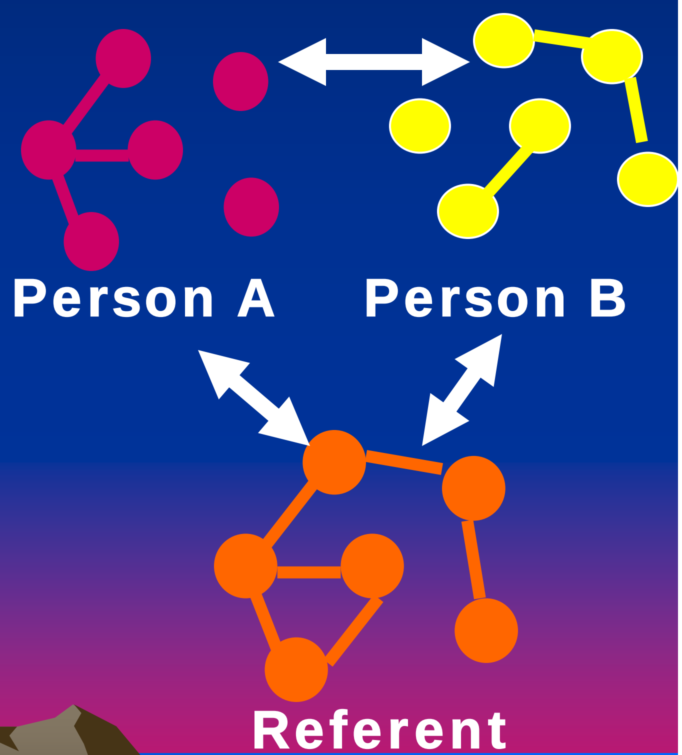
- Measures tend to assume homogeneous groups
- Measures tend to target collective level
- Aggregation methods are limited
- Measures are needed that target the more dynamic and fleeting knowledge
- Measures are needed that target different types of long-term collaborative knowledge
- A broader range of knowledge elicitation methods is needed
- A need for streamlined and embedded measures
- Newly developed measures require validation

# Assessing Collaborative Performance and Cognition

# The "*Apples and Oranges*" *Problem*



Measures to assess collaborative knowledge often assume knowledge homogeneity among group members.



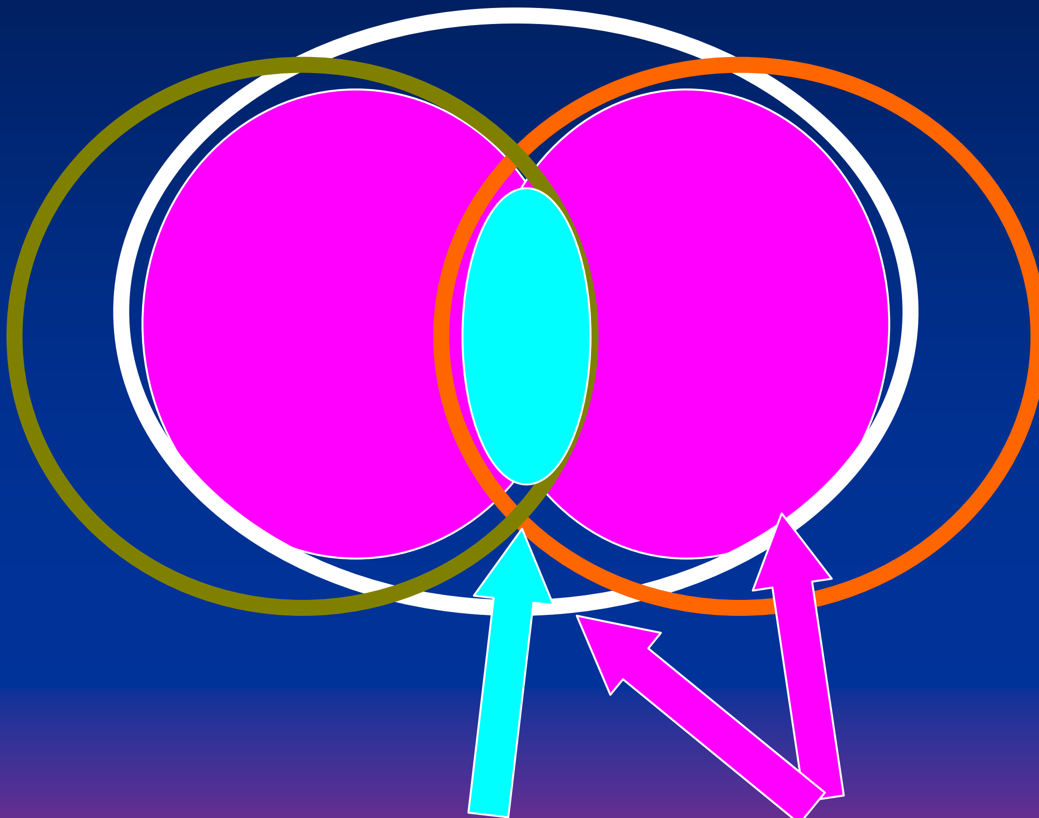

- Shared knowledge = similar knowledge
- Accuracy is relative to single referent

# The Groups We Study Consist of "Apples and Oranges"

Airport Incident Command Center


Telemedicine

# “Shared” Knowledge



**Shared = Common and Complementary**

# An Approach to the *Apples and Oranges* Problem

***Measures of team knowledge with heterogeneous accuracy metrics***

# Experimental Context

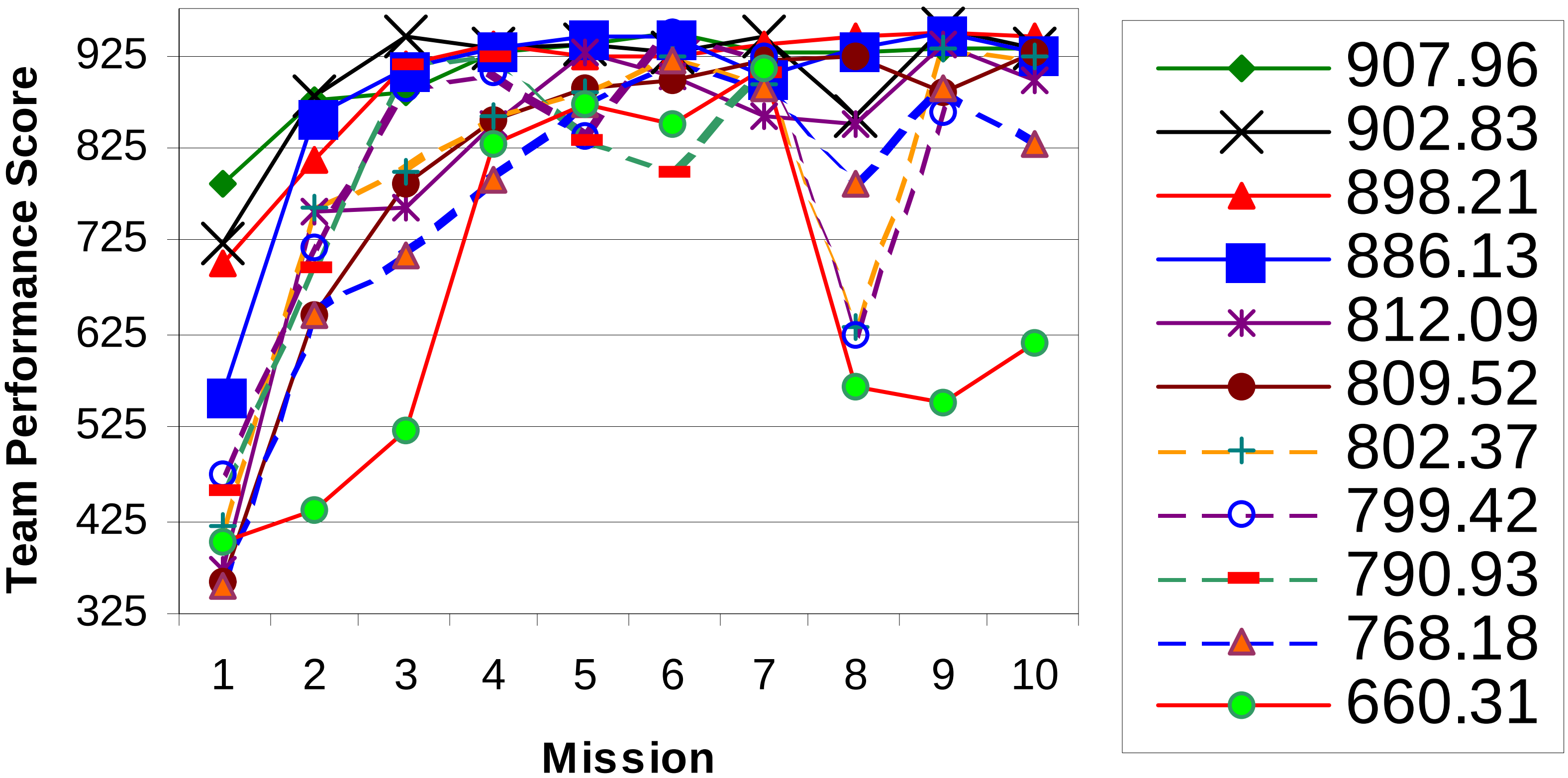
- **Five studies:** Two different 3-person tasks: UAV (Uninhabited Air Vehicle) and Navy helicopter rescue-and-relief
- **Procedure:** Training, several missions, knowledge measurement sessions
- **Manipulate:** co-located vs. distributed environments, training regime, knowledge sharing capabilities, workload

# Experimental Context

## MEASURES

- Team performance: composite measure
- Team process: observer ratings and critical incident checklist
- Other: Communication (flow and audio records), video, computer events, leadership, demographic questions, working memory, situation awareness
- Taskwork & Teamwork Knowledge

## Scores from completed missions for all teams



# Long-term Taskwork Knowledge

## ❖Factual Tests

The camera settings are determined by a) altitude, b) airspeed, c) light conditions, d) all of the above.

## ❖Psychological scaling

How related is *airspeed* to *restricted operating zone*?

# Long-term Teamwork Knowledge

Given a specific task scenario, who passes what information to whom?

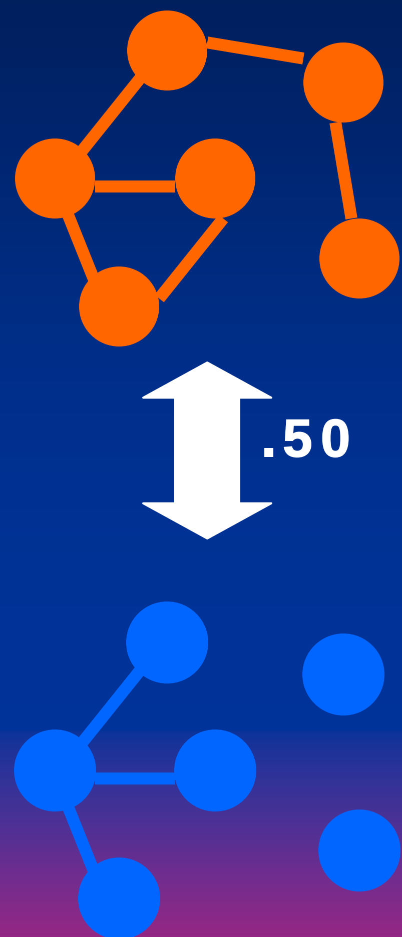
## Teamwork Checklist

___AVO gives airspeed info to PLO
___DEMPC gives waypoint restrictions to AVO
___PLO gives current position to AVO

**AVO= Air Vehicle Operator**
**PLO = Payload Operator**
**DEMPC = Navigator**

# Traditional Accuracy Metrics

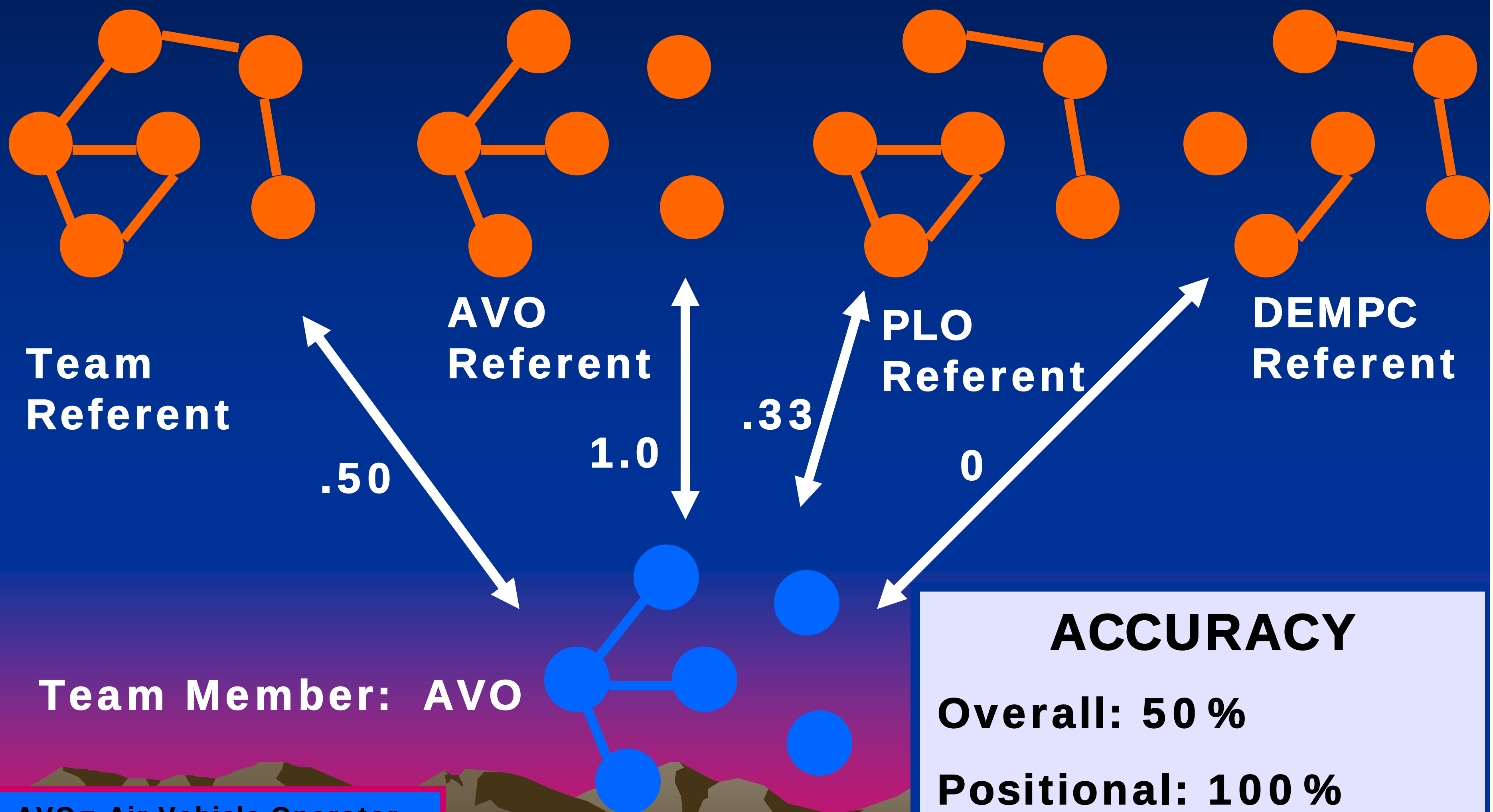
**Team Referent**




**Team Member:**
**Air Vehicle Operator**

**50% ACCURACY**

# Heterogeneous Accuracy Metrics




**ACCURACY**

**Overall: 50%**

**Positional: 100%**

**Interpositional: 17%**

AVO= Air Vehicle Operator
PLO = Payload Operator
DEMPC = Navigator

# Results Across Studies

- Taskwork knowledge is predictive of team performance

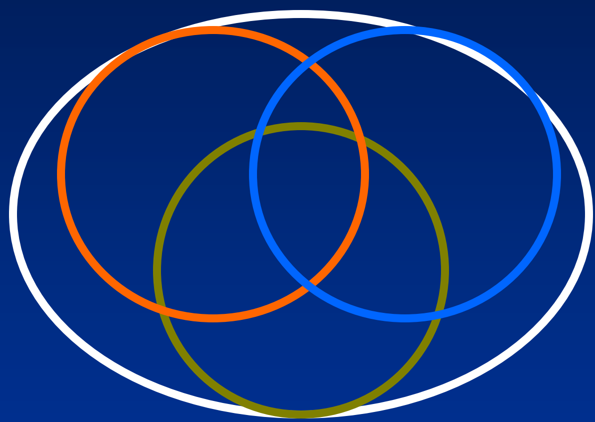
  But...
- True for psychological scaling, not factual tests
- Timing of knowledge test is critical

# Knowledge Profiles of Two Tasks

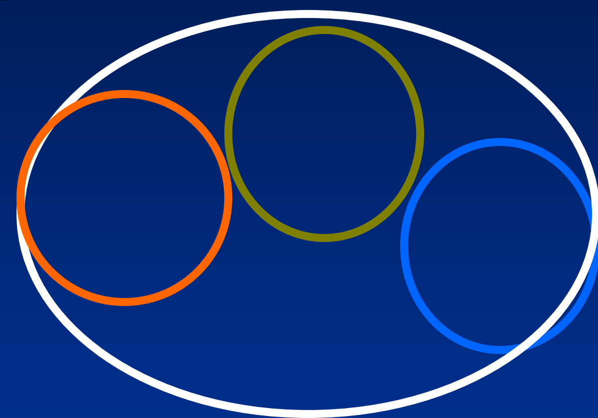
**Knowledge profile characterizing effective teams depends on task (UAV vs. Navy)**

| | Knowledge Profile | |
|---|---|---|
| Knowledge metric | Common (UAV) | Distributed (Navy helicopter) |
| Overall accuracy | + | 0 |
| Intrateam similarity | + | 0 |
| Positional accuracy | + | + |
| Interposit. accuracy | + | 0 |

# Knowledge Profiles of Two Tasks



Common



Complementary

| ***UAV Task*** | ***Navy Helicopter Task*** |
|---|---|
| **Command-and-Control** | **Planning and Execution** |
| **Interdependent** | **Less interdependent** |
| **Knowledge sharing** | **Face-to-Face** |

# A Cross-Training Study in Retrospect

- Examined effects cross-training vs. other training regimes on collaborative performance and cognition
- Unlike previous studies, cross-training had no performance benefit
- Cross-training, did increase interpositional taskwork and teamwork knowledge
- Perhaps knowledge profile for that task (specialization) was at odds with cross-training
- Demonstrates benefits of assessing collaborative cognition

# However…

## The descriptive information associated with cognitive assessment is not sufficiently diagnostic

### Symptoms vs. Diagnoses?

- Expert chess players can remember many more meaningful chess positions than chess novices
- Experienced fighter pilots and undergraduate pilot trainees organize flight maneuver concepts differently
- Good UAV teams have interpositional knowledge
- Effective teams communicate less than ineffective ones

# Toward Diagnosis of Collaborative Performance

# To Move From Assessment to Diagnosis ….

Need to connect clusters of symptoms to diagnosis of team dysfunction or excellence

For example…

- Inefficient communication and low teamwork knowledge → poor team situation awareness
- Poor positional taskwork knowledge and coordination failures → faulty mental model

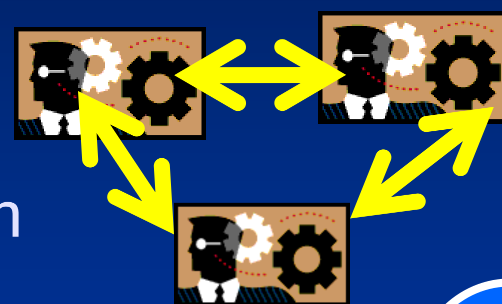
# To Move From Assessment to Diagnosis ….

Also, in operational environments diagnosis is valuable to the extent that it is...

- **Conducted in real time with task performance (i.e., task-embedded, automated measures)**
- **Or better yet …prior to task performance (based on performance precursors)**
  - Leading indicator
  - Resident pathogens
  - Non-routine events

# Communication as a Window to Collaborative Cognition

- Observable; Think aloud “in the wild”
- Reflects collaborative cognition at the holistic level; is collaborative cognition
- Embedded in the task
- But...labor intensive transcription, coding, and interpretation
- Exploit its richness by automating analyses



# Our Approach to Communication Analysis

- Communication Flow Analysis
- Content Analysis Using Latent Semantic Analysis

# Analyzing Flow: CERTT Lab ComLog Data


**Team members use push-to-talk intercom buttons to communicate. At regular intervals speaker and listener identities are logged**

# Analyzing Flow: ProNet--Procedural Networks

- Nodes define events that occur in a sequence
- An Example from UAV study: 6 nodes: Abeg, Aend, Pbeg, Pend, Dbeg, Dend
- ProNet: Find representative event sequences

**Quantitative: Chain lengths-->Performance**

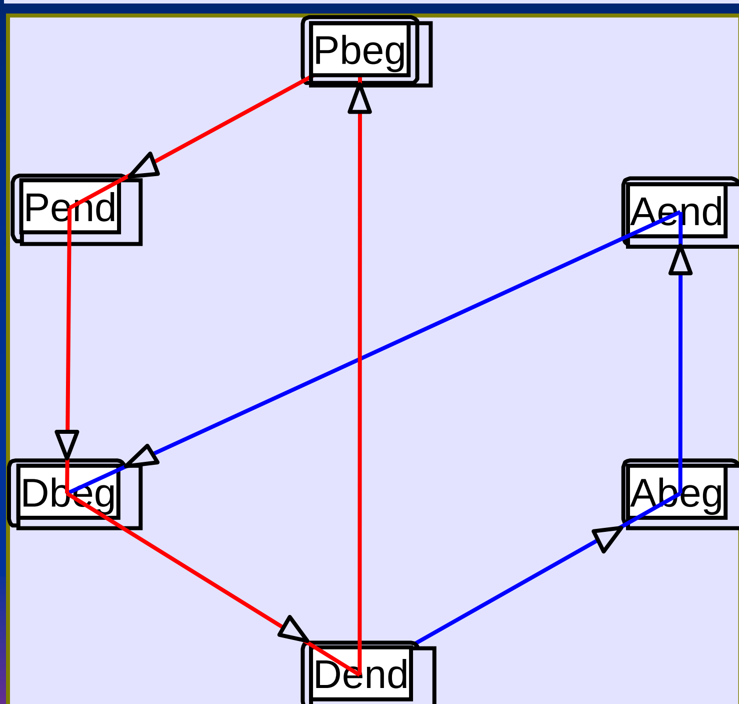
Mission 2: $R^2$ = .509, $F(2, 8)$ = 4.144, $p$ = .058

Mission 3: $R^2$ = .275, $F(1, 9)$ = 3.415, $p$ = .098

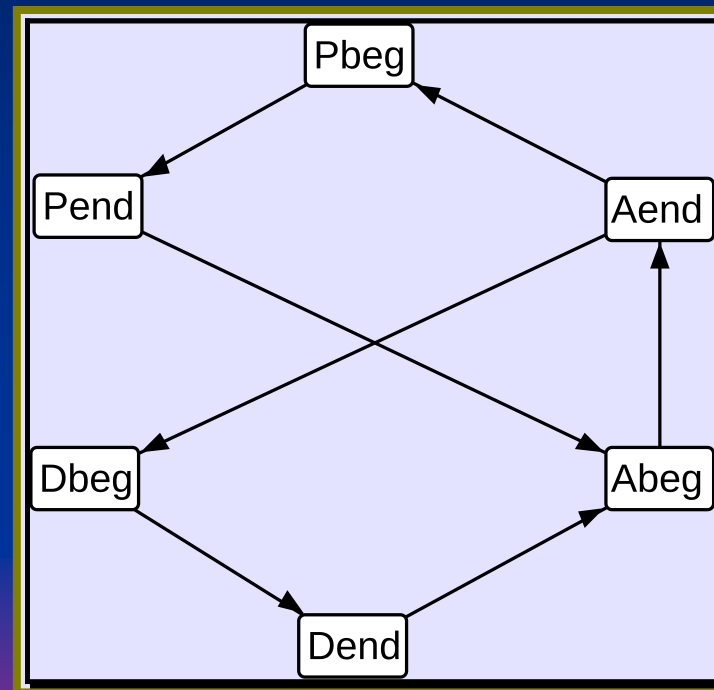
Mission 5: $R^2$ = .628, $F(2, 8)$ = 5.074, $p$ = .051

# Analyzing Flow: ProNet-- Procedural Networks

**Qualitative: Communication patterns predictive of performance**




Team 2 before PLO-DEMPC's fight




Team 2 after PLO-DEMPC's fight

AVO= Air Vehicle Operator
PLO = Payload Operator
DEMPC = Navigator

# Analyzing Flow: Other Approaches

- Measure of speaker dominance
- Deviations from ideal flow
- Clustering model-based patterns

# Content Analysis with Latent Semantic Analysis (LSA)

**Landauer, Foltz, & Laham, 1998**

- A tool for measuring cognitive artifacts based on semantic information.
- Provides measures of the semantic relatedness, quality, and quantity of information contained in discourse.
- Automatic and fast.
- We can derive the meaning of words through analyses of large corpora.

# Content Analysis with Latent Semantic Analysis (LSA) (continued)

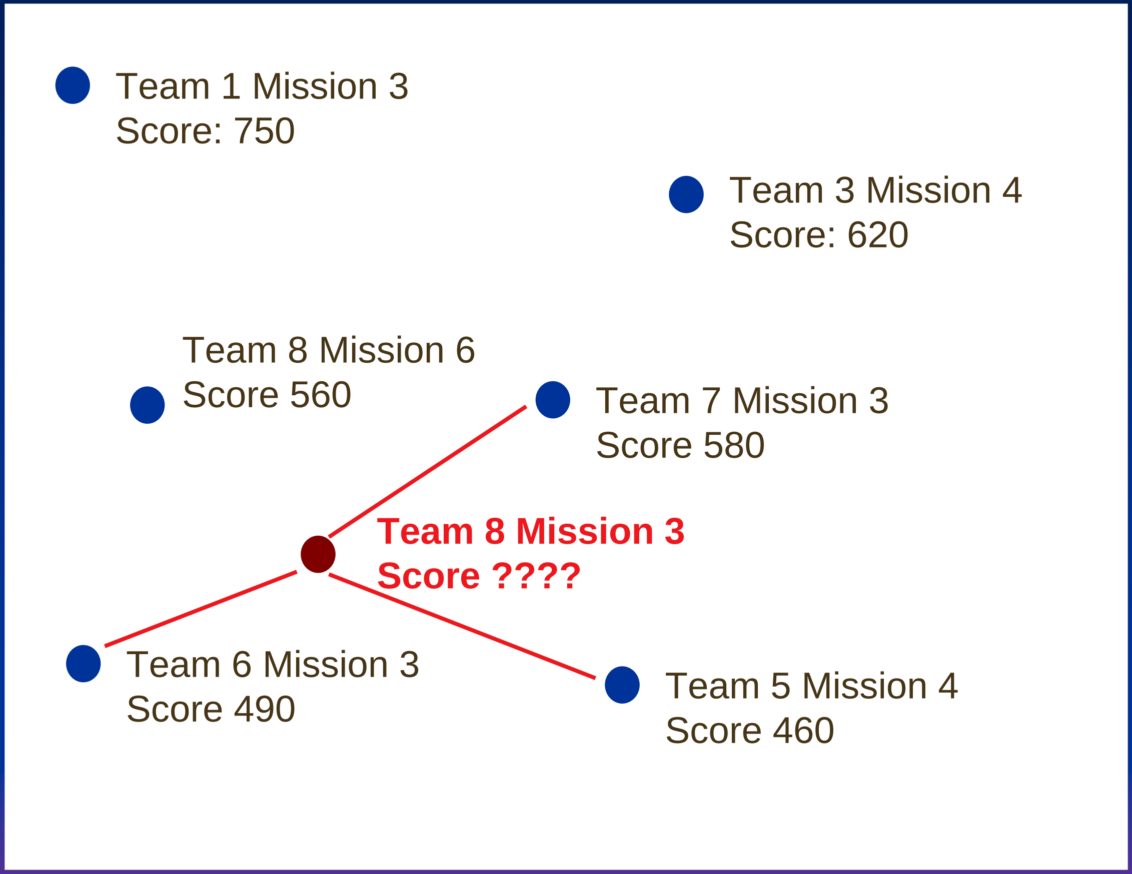
- Large constraint satisfaction problem of estimating the meaning of many passages based on their contained words (like factor analysis)
- Method represents units of text  (words, sentences, discourse, essays) as vectors in a high dimensional semantic space based on correlations of usage across text contexts
- Compute degree of semantic similarity between any two units of text

# Content Analysis with Latent Semantic Analysis (LSA)

## An Example from UAV Study 1

- 67 Transcripts from missions 1-7
  - XML tagged with speaker and listener information
  - ~2700 minutes of spoken dialogue
  - 20,545 separate utterances (turns)
  - 232,000 words (660 k bytes of text)
- Semantic Space: 22,802 documents
  - Utterances from dialogues
  - Training material
  - Interviews with domain experts
- Derived several statistical measures of the quality of each transcript

# Content Analysis with Latent Semantic Analysis (LSA)




**LSA-based communication score predicts performance (r =.79).**

# Other Significant Variables

- Variance of scores of similar dialogues r=-.58
- Vector length, r=-.35
- Number of words, r=-.34
- Zipf $R^2$, r=-.47
- Percent non-function words, r=.34
- ....
- Five factor RMMR model: r=.76

**Conclusion: We can accurately predict team performance from dialogues as a whole.**

# Analyzing Content: Other Approaches

- Automatic transcript coding
- Coherence in team dialogue
- Measures of individual contributions

# Implications of Communication Work

- Collaborative cognition is revealed through discourse
- Measurement of the team as a whole, as well as individuals
- Techniques move toward automated analyses of the content of team dialogues
- Avoids tedious hand coding, keeps high reliability
- Automation will allow for task-embedded measures that assess and diagnose in real-time

# Conclusions

# Summary

- Understanding collaborative cognition is critical for diagnosis of team dysfunction or excellence and later intervention
- Assessment is only a first step
- Diagnostic information is needed
- In operational environments diagnosis needs to be task-embedded, automatic, and forward-looking

# Implications & Applications

- Suggestions for aiding collaborative cognition through training or technology.
- Selecting/composing teams for optimal collaboration
- On-line monitoring of collaborative cognition in high-risk environments

# Questions or Comments?

# Contact
# Nancy J. Cooke
# Arizona State University

ncooke@asu.edu

http://psych.nmsu.edu/CERTT/